

Description

SIMPLIFIED TOP OXIDE LATE PROCESS

BACKGROUND OF INVENTION

[0001] The present invention generally relates to integrated circuit (IC) memory devices and, more particularly, to a process for protecting thin gate oxides by covering polysilicon studs during gate oxidation in vertical metal oxide semiconductor field effect transistor (MOSFET) dynamic random access memory (DRAM) arrays.

[0002] A DRAM circuit usually includes an array of memory cells interconnected by rows and columns, which are known as wordlines (WLs) and bitlines (BLs), respectively. Typically, a DRAM memory cell comprises a MOSFET connected to a capacitor, such as a vertical MOSFET having a trench capacitor. Vertical MOSFETs allow for an effective size reduction in bit densities, and as such, the use of such vertical MOSFETs has increased over the years as the need to overcome the scalability limitations of planar MOSFET DRAM access transistors has also increased.

[0003] The use of vertical MOSFETs is not yet widespread and

several characteristics need to be optimized. In particular, in the formation of a memory cell in vertical MOSFET DRAM arrays, an array top oxide (ATO) is needed to isolate the passing word-lines from active areas on the substrate. Typical processes for forming these isolating ATO areas include forming an array top oxide layer in the presence of a pad nitride layer of the DRAM. A variety of methods exist in the art for forming array top oxides, such as, top oxide early, top oxide nitride and top oxide late procedures.

[0004] Typically, in a top oxide early (TOE) procedure, an array top oxide area is formed with pad nitride being present in the support. Any remaining pad nitride in the support is then removed, a gate oxide is grown and then a gate poly is deposited. Similarly, in top oxide nitride (TON) procedures, an array top oxide area is formed after the pad nitride has been stripped in both the array and support areas. Array top oxide is then removed in the support followed by gate oxidation and deposition of a gate poly. However, a disadvantage of such procedures is that both TOE and TON processing require a large number of masks, and thus processing steps, for fabrication of these ATO areas.

[0005] Further procedures for forming array top oxide areas include a top oxide late (TOL) scheme. In TOL processing, a pad nitride layer is removed from the entire surface of the DRAM. Thereafter, a gate oxide is grown and a gate poly is deposited in the support area. An array top oxide layer is then deposited after removal of the gate poly in the array. The advantages of TOL processing include reduced number of masks, processing steps and manufacturing costs as well as providing a litho-friendly structure with a flat surface after the top oxide CMP in comparison to TOE and TON processes. However, it is significantly more difficult to achieve compared to TOE and TON.

[0006] Additionally, during TOL fabrication of the gate oxidation for support devices in vertically-arranged-device memory cells arsenic from the gate poly of the vertical array devices, i.e., polysilicon stud, can out-diffuse and react with the nitrogen. This undesirably results in thicker oxides in the thin gate oxide regions, which in turn, increases the V_t of the support device. Numerous steps are required, including depositing a liner layer over the polysilicon stud, in order to prevent the out-diffusion of arsenic during the formation of the gate oxidation for support devices in vertical DRAM.

[0007] Accordingly, a need continues to exist in the art for providing further improved, simplified techniques of forming an array top oxide area in vertical DRAM that require a reduced number of masks, processing steps and manufacturing costs, provides a litho-friendly structure, and prevents out-diffusion of arsenic from the gate poly during TOL fabrication of the gate oxidation for the support devices in the vertical memory cell.

SUMMARY OF INVENTION

[0008] Bearing in mind the problems and deficiencies of the prior art, it is therefore an object of the present invention to provide an improved and simplified method for forming an array top oxide area in vertical DRAM.

[0009] Another object of the present invention is to provide a method for forming an array top oxide area in vertical DRAM via an improved, simplified top oxide late (TOL) scheme that protects thin gate oxides during gate oxidation by covering the polysilicon studs of the vertical DRAM.

[0010] It is another object of the present invention to provide a method for forming an array top oxide area in vertical DRAM that requires a reduced number of masks and processing steps.

- [0011] A further object of the invention is to provide a method for forming an array top oxide area in vertical DRAM that prevents out-diffusion of arsenic from the gate poly during TOL fabrication of the gate oxidation for the support devices in the vertical memory cell.
- [0012] It is yet another object of the present invention to provide a method for forming an array top oxide area in vertical DRAM that reduces costs.
- [0013] Still other objects and advantages of the invention will in part be obvious and will in part be apparent from the specification.
- [0014] The above and other objects and advantages, which will be apparent to one of skill in the art, are achieved in the present invention, which, is directed to in a first aspect a method of providing an array top oxide over an array of trenches containing trench capacitors and overlying vertical transistors. The method includes providing a substrate having a first area and a second area. Preferably, the first area is an array area, while the second area is a support area. A sacrificial oxide layer is provided over the first and second areas and such sacrificial oxide layer is removed only from the second area. A gate oxide layer is then provided over the second area and a gate conductor layer de-

posited over both first and second areas. The gate conductor layer is then removed only from the first area. Subsequently, an array top oxide layer is deposited over the first and second areas, and the array top oxide layer is removed from the second area so as to leave remaining portions of the array top oxide layer only in the first area.

[0015] Prior to the step of growing the sacrificial oxide layer, the method may further include removing a pad nitride layer in both the first and second areas to expose a pad oxide layer in both the first and second areas, removing at least a portion of the pad oxide layer from both the first and second areas and then growing the sacrificial oxide layer over exposed surface areas of the substrate. The first and second areas may then be implanted.

[0016] The sacrificial oxide layer may be removed from the second area only by providing a mask over the first area to protect the first area during the sacrificial oxide layer removal processing. This mask is removed from second area prior to growing the gate oxide.

[0017] Gate oxide is grown at least in the second area, and gate conductor is deposited in the first and the second area. Further, the gate conductor layer may be removed from the first area also by providing a mask over the second

area to protect the second area during the gate conductor layer removal processing. This mask is removed prior to depositing the array top oxide layer.

[0018] In another aspect, the invention relates to a method of providing an array top oxide over an array of trenches containing trench capacitors and overlying vertical transistors. The method includes providing a substrate having an array area and a support area and then providing a sacrificial oxide layer over the array and support areas. The sacrificial oxide layer is removed from the support area and then a gate oxide layer is provided over the support area. A gate conductor layer is then deposited over the array and support areas and the gate conductor layer removed from the array area only. An array top oxide layer is then deposited over the array and support areas. This array top oxide layer may be deposited by high density plasma deposition. The array top oxide layer is removed from the support area by planarization so as to leave remaining portions of the array top oxide layer only in the array area. In removing the sacrificial oxide layer only from the support area, a first mask is provided over the array area to protect such array area during the sacrificial oxide layer removal processing. This first mask is

then removed from the array area prior to providing the gate oxide layer over the support area. A second mask is then provided over the support area to protect the support area during the gate conductor layer removal processing. The second mask is also removed prior to depositing the array top oxide layer. These first and second masks may both comprise etch support lithography masks.

[0019] In accordance with the invention, the step of removing the sacrificial oxide layer only from the support area prevents out-diffusion from the array area, particularly arsenic out-diffusion, while during the step of growing the gate oxide layer over the support area, the sacrificial oxide layer protects the underlying array area to also prevent any out-diffusion there-from, particularly arsenic out-diffusion.

BRIEF DESCRIPTION OF DRAWINGS

[0020] The features of the invention believed to be novel and the elements characteristic of the invention are set forth with particularity in the appended claims. The figures are for illustration purposes only and are not drawn to scale. The invention itself, however, both as to organization and method of operation, may best be understood by reference to the detailed description which follows taken in

conjunction with the accompanying drawings in which:

- [0021] Fig. 1 is a cross sectional view of a DRAM for processing in accordance with the invention.
- [0022] Fig. 2 is a top plan view of the DRAM of Fig. 1.
- [0023] Fig. 3 is a cross sectional view of the step of deglazing the DRAM of Figs. 1 and 2.
- [0024] Fig. 4 is a cross sectional view of the step of stripping a pad oxide layer from array and support areas of the DRAM of Fig. 3, followed by growing a sacrificial oxide layer over all exposed surface areas of such stripped DRAM.
- [0025] Fig. 5 is a cross sectional showing the step of stripping the sacrificial oxide of Fig. 4 in the support area only using a block mask 80 that entirely covers only the array area 102.
- [0026] Fig. 6 is a top plan view of the process step of Fig. 5.
- [0027] Fig. 7 is a cross sectional view of the step of removing the block mask, followed by growing a gate oxide layer at least over support areas of the DRAM of Figs. 5 and 6.
- [0028] Fig. 8 is a cross sectional view of the resultant DRAM after further processing of the pre-finished DRAM of Fig. 7 including the steps of depositing a gate conductor material over the surface of the substrate, removing the gate con-

ductor material from the array area only using a second mask, removing the second mask, depositing an array top oxide layer over the substrate surface, and then planarizing the substrate surface to leave the array top oxide only in the array area and expose conductive poly of vertical transistors within array areas of the resultant DRAM.

DETAILED DESCRIPTION

- [0029] In describing the preferred embodiment of the present invention, reference will be made herein to Figs. 1–8 of the drawings in which like numerals refer to like features of the invention. Features of the invention are not necessarily shown to scale in the drawings.
- [0030] Referring to Fig. 1, there is shown a cross-sectional view of a section of a DRAM 100 containing vertical MOSFET arrays formed in a semiconductor wafer substrate 10 along cross sections X1–X2 and X2–X3. For ease of understanding, Fig. 2 illustrates a top down view of the plurality of trench locations X1, X2 and X3 of DRAM 100. The DRAM 100 is provided with both an array area 102 of the substrate 10 and a support area or circuitry 104 of the substrate 10.
- [0031] Initially, pad oxide layer 22 and pad nitride layer 40 are deposited over the substrate and then a plurality of deep

trenches 20 are formed from the top surface down into the substrate 10, preferably a silicon substrate, by known patterning and etching techniques. These deep trenches 20 are formed within the DRAM 100 array area 102 of the substrate 10. An outer buried plate in the substrate and a node dielectric (not shown) are then formed. A collar oxide 24 is formed within a middle portion of each of these deep vertical trenches 20 for preventing the formation of a parasitic transistor and then the trenches filled with polysilicon to a sufficient depth within the lower portion of the deep trenches to act as the conductor for the storage node.

[0032] A trench top oxide layer 22" is deposited within the deep trenches and then a gate oxide is deposited in the upper portions 20a of the trenches over the trench top oxide 22". Subsequently, these upper portions of the deep trenches are filled with additional polysilicon to a sufficient depth to serve as the gate conductors (GC) for the DRAM cell. The DRAM is provided with vertical sidewall spacers 52 within the deep trenches 20, whereby these sidewall spacers are formed by depositing a SiN layer and etching using an anisotropic dry etch technique, such as, Reactive Ion Etch (RIE). The remaining portions of the deep

trenches are completely filled with additional polysilicon and the substrate surface area is then planarized to remove any excess polysilicon fill material.

[0033] Thus, as shown in Fig. 1, the trench top oxide layer 22" divides the deep trenches into a vertical transistor portion residing in the upper portion of the trench, above the trench top oxide 22", and a trench capacitor portion residing in the lower portion of the trench, below the trench top oxide 22". The trench top oxide layer isolates the pass transistor from the trench capacitor by isolating the gate oxide in the upper portion of the trenches from the storage node in the lower portion of the trenches. As shown, the collar oxide 24 is located within each of the trenches 20 to a depth just below the trench top oxide 22", i.e., below the transistor portion of deep trenches 20, so as to prevent outward leakage of charge stored within the capacitor (not shown) located in the lower portion of trenches 20.

[0034] The DRAM 100 further includes a buried strap poly 28 that out-diffuses into the substrate to form the buried strap within the lower portion of the trenches. A region of the vertical pass transistor 60 is formed in the substrate by known techniques. The DRAM 100 also includes gate

26 within the upper portions of the trenches 20, which, is the gate of the pass transistor.

[0035] Shallow trench isolations are formed in both the array area 102 and support area 104 of DRAM 100 by patterning and etching in both the array 102 and support 104 areas, which, are subsequently filled with dielectric, preferably an oxide fill. These isolation trenches 30 are adjacent to each other and isolate adjacent deep trenches 20 within separate array areas as shown by the top plan view of DRAM 100 in Fig. 2, which, is a top down view of the array area 102 of DRAM 100 as shown in Fig. 1 along the directional X1, X2, and X3. Any excess dielectric material is then removed, preferably, by a chemical mechanical polishing technique.

[0036] The above steps of processing DRAM 100 proceed in a normal fashion through the steps of forming the vertical deep trenches 20, having gate conductors and capacitors therein, through planarizing the substrate surface, such as by chemical mechanical polishing, after formation of isolation trenches 30.

[0037] Referring to Fig. 3, the entire substrate surface of DRAM 100 is then deglazed, preferably, by a wet etching. The deglazing step is crucial as it removes any residual oxide

residing over the pad nitride layer 40 in both the array and support areas of the DRAM, as well as slightly recesses the isolation trenches to a desired height. Subsequently, after any residual oxide is stripped and isolation trenches 30 recessed, the pad nitride layer 40 is stripped in both array areas 102 and support areas 104 of substrate 10, thereby leaving the pad oxide layer 22 exposed as shown in Fig. 3.

[0038] Referring to Fig. 4, the pad oxide layer 22 is then stripped from both the array areas 102 and the support areas 104 on substrate 10. Preferably, the pad oxide 22 is stripped using a wet etch technique. A sacrificial oxide layer 70 is then grown over all exposed surface areas of the DRAM, in both array and support areas, for the array and the support implants. However in the preferred embodiment, wherein the isolation trenches 30 comprise oxide, grown oxide will not form over the oxide isolation trenches 30 as shown in Fig. 4.

[0039] After the sacrificial oxide layer 70 is grown over all exposed surfaces in both array and support areas, a block mask 80 is provided only over the array areas 102, as shown in the cross sectional view of Fig. 5 and the top plan view of Fig. 6. This block mask 80 preferably com-

prises an etch support lithography mask that is spin-coated so as to entirely cover only the array area 102, while not covering any of the support area 104.

[0040] Using the block mask 80 that entirely covers the array area only, the sacrificial oxide layer 70 is completely stripped from the support area only. During this process, the block mask 80 protects the entire underlying array area of the substrate. Conventional wet etch chemistries, such as buffered hydrofluoric acids, are preferably used to strip the sacrificial oxide in the support area only using block mask 80 to protect the array area.

[0041] In accordance with the invention, the array area is protected during removal of the sacrificial oxide layer 70 in the support area thereby preventing any out-diffusion of arsenic from the array area of the DRAM during gate oxidation. The present invention advantageously provides processing ease as the block mask 80 is directly deposited over the grown sacrificial oxide layer 70, only in the array area, without the use of any additional processing layers. That is, the present invention significantly simplifies the process of forming an array top oxide in the array while protecting thin gate oxides by covering polysilicon studs during gate oxidation.

[0042] As shown in Fig. 7, after the sacrificial oxide is removed from the support area only using block mask 80, the block mask 80 is removed from the surface of the substrate. This is preferably accomplished by stripping block mask 80 by a fusion strip technique. Once the block mask 80 is stripped, a gate oxide layer(s) 75 is thermally grown only over exposed substrate 10 surface areas in the support area 104 of the substrate. During this step, the array area is already covered with grown oxide layer 70, and as such, an insignificant amount, if any, of gate oxidation will grow over the oxide layer 70 in the array area.

[0043] A critical feature of the invention is that during this step of growing gate oxide layer 75, any arsenic out-diffusion of the As-doped polysilicon studs in the array area is prevented as such studs are entirely covered with grown oxide layer 70. Without this grown oxide layer 70 covering the polysilicon studs, arsenic can out-diffuse and react with nitrogen that has been implanted in thin gate oxide regions to grow thinner oxides, and hence, results in an increase of the V_t 's for thin gate oxide devices.

[0044] Referring to Fig. 8, after the gate oxidation is formed in the support area, a gate conductor material 90 is deposited over the entire substrate surface, directly over

gate oxide 75, and then using a second block mask that covers only the support area 104, the gate conductor material is entirely removed from the array area only. In so doing, the second block mask protects the support area during the process of removal of the gate conductor from the array area. This second block mask is then removed from the substrate surface.

[0045] Subsequently, an array top oxide layer 80 is deposited over both array 102 and support 104 areas, preferably via high-density plasma deposition. Once the array top oxide is deposited, the substrate is planarized, preferably by chemical mechanical polishing, so as to leave the array top oxide layer 80 only in the array area 102, therein resulting in the structure as depicted in Fig. 8. In planarizing the substrate surface, planarization preferably stops once the top surface of trenches 20 are reached, therein removing the oxide layer 70 from the trenches 20, exposing a surface of the trenches 20 and leaving remaining portions of gate conductor material 90 only in the support regions of the substrate.

[0046] The DRAM may be further processed to form gate conductors (wordlines), bitlines, and back end interconnects by conventional techniques. In so doing, the exposed poly

26 is used as the base of a gate/wordline structure for forming the gate contact to the vertical transistors 60, which extends perpendicular to the plane of the substrate to form at least one of the wordlines of the array. A second wordline, referred to as the "passing wordline," is formed between the two adjacent poly structures 26. This passing wordline makes contact to DRAM cells behind and in front of the trenches 20.

[0047] Thus, in accordance with the invention, a simplified process flow, preferably a simplified TOL process, is provided for providing an array top oxide over an array of trenches containing trench capacitors and overlying vertical transistors. As demonstrated in the below comparative table, conventional processes for providing array top oxides require numerous processing steps, such as, depositing additional liner layers over the sacrificial oxidation layer for the purpose of preventing out-diffusion from the array area, whereas the present invention eliminates the need for any additional processing layers by directly forming a blocking mask over the sacrificial oxidation layer and using such mask to remove the sacrificial oxidation layer from the support region only.

[0048]

TABLE I

Prior Art Process Flow	Process Flow of the Invention
Pad SiN strip	Pad SiN strip
Pad Oxide strip (pullback)	Pad Oxide strip (pullback)
Sacrificial oxidation	Sacrificial oxidation
Nitride Liner	Implants
ES Mask/Nitride RIE/Resist strip	(Optical Nitride spacer)
Support Implants	ES Mask/Sac Ox wet strip/Resist strip
Sacrificial Oxide strip	Gate oxidation
Gate oxidation	

[0049] Advantageously, the present invention provides an improved TOL process with a partial lithography-assisted sacrificial oxide strip to prevent arsenic out-diffusion from polysilicon studs during gate oxidation. In particular, it prevents arsenic out-diffusion during gate oxidation from polysilicon studs by completely covering such studs with an oxide layer, thereby maintaining nitrogen amounts in the thin gate oxides region, and hence, maintaining gate oxide thickness and avoiding any increase in V_t 's for thin gate devices. The invention eliminates the need for additional liner layers over the sacrificial oxidation layer for the purpose of preventing out-diffusion from the array area by directly forming a blocking mask over the sacrificial oxidation layer and using such mask to remove the sacrificial oxidation layer from the support re-

gion only.

[0050] While the present invention has been particularly described, in conjunction with a specific preferred embodiment, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art in light of the foregoing description. It is therefore contemplated that the appended claims will embrace any such alternatives, modifications and variations as falling within the true scope and spirit of the present invention.